

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 06-222797

(43)Date of publication of application : 12.08.1994

(51)Int.Cl.

G10L 9/14
G10L 9/18

(21)Application number : 05-008737

(71)Applicant : NEC CORP

(22)Date of filing : 22.01.1993

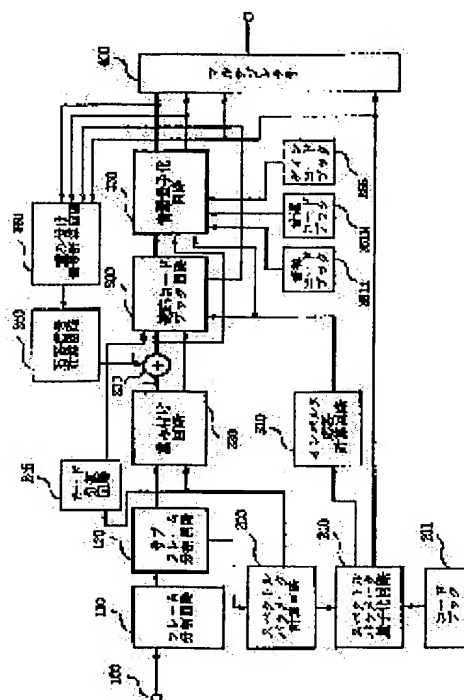
(72)Inventor : OZAWA KAZUNORI

(54) VOICE ENCODING SYSTEM

(57)Abstract:

PURPOSE: To provide an voice encoding system obtaining high tone quality with 4.8kb/s bit rate or below.

CONSTITUTION: This system is provided with a spectrum parameter calculation part 200 dividing an voice signal to frames, dividing it to subframes and calculating the spectrum parameter of the subframe, a specturum parameter quantizing part 210 quantizing the spectrum parameter, a weighting part 230 calculating a characteristic amount from the voice signal and performing acoustic weighting to the voice signal by using the mode grouping part 245 of the frame voice signal and the spectrum parameter, an adaptive code book part 300 using a mode grouping result and the quantized spectrum parameter and the voice signal, and obtaining a prameter showing a pitch period and a sound source quantizing part 350 using the weighted signal and the output of the adaptive code book part 300 and the spectrum parameter and the quantized spectrum parameter, and retrieving sound source code books 3511-351N consisting of plural stages and a gain code book 355, and quantizing the source signal.



LEGAL STATUS

[Date of request for examination]

22.01.1993

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application]

converted registration]

[Date of final disposal for application]

[Patent number] 2746039

[Date of registration] 13.02.1998

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates a sound signal to a low bit rate and the voice coding method for encoding in high quality especially with the bit rate of 4.8 or less kb/s.

[0002]

[Description of the Prior Art] As a method which encodes a sound signal with the low bit rate of 4.8 or less kb/s, conventionally For example, Schrader (M. Schroeder) and the "code-EKISAIEDO linear PUREDI cushion by Atar (B. Atal): High quality speech ATO Berry Low bit REITSU () [Code-excited] linear prediction: High quality speech at very low bit rates", Sound, voice, and the minutes of the international congress about signal processing (Proc.ICASSP), In 1985 A paper (937 thru/or 940 pages) (reference 1), "improved speech quality - and - EFISHIENTO vector KUONTAIZEISHON Inn S IE rupee () by clay gin (Kleijn) [Improves speech quality] and efficient vector quantization in SELP", The CELP (Code Excited LPC Coding) method indicated by 155 thru/or the 158-page paper (reference 2) is learned in the minutes (Proc.ICASSP) of the international congress about sound, voice, and signal processing, and 1988. The spectrum parameter which carries out linear prediction (LPC) analysis of the sound signal for every (for example, 20ms) frame, and expresses the spectral characteristics of a sound signal with this method by the transmitting side is extracted. Divide a frame into a subframe (for example, 5ms) further, and the parameter (a delay parameter and gain parameter) in an adaptation code book is extracted based on the past excitation signal for every subframe. The optimal sound-source code vector is chosen from the sound-source code book (vector quantization code book) which consists of a noise signal of the class defined beforehand to the remainder signal which carried out pitch prediction and which carried out pitch prediction of the sound signal of a subframe with the adaptation code book, and was searched for, and the optimal gain is calculated. Selection of the optimal sound-source code vector is performed so that the error power of the signal compounded with the selected noise signal and the above-mentioned remainder signal may be minimized. And the parameter extracted from the index showing the class of selected sound-source code vector, the optimal gain, the above-mentioned spectrum parameter, and the adaptation code book is transmitted. Explanation of a receiving side is omitted.

[0003]

[Problem(s) to be Solved by the Invention] By the conventional method of the reference 1 and 2 mentioned above, in order to acquire good tone quality, there was need that the size of a sound-source code book is sufficiently large (for example, 10 bits). For this reason, the huge amount of operations was needed for retrieval of a sound-source code book. The memory space furthermore needed was also huge (for example, the case of 10 bits 40 dimensions 40 K-word memory space), and it was difficult to realize hardware in a compact. Moreover, in order to reduce a bit rate, when it increases frame length and subframe length and a number of dimension is increased, without reducing the number of bits of a sound-source code book, the amount of operations has the trouble of increasing very notably.

[0004] As an approach of reducing the size of a code book For example, "multiple stage vector KUONTAIZEISHON forehand speech coding by JUANGU and others (B. Juang) () [Multiple Stage vector quantization] for speech As indicated by 597 thru/or the 600-page paper (reference 3) in coding", the minutes (Proc.ICASSP) of the international congress about sound, voice, and signal processing, and 1982 A code book is divided and constituted in multistage and the multistage vector-quantizing method for looking for each code book independently is learned. L step also in the whole is also compared with one step of B bit, and the amount of operations which code book retrieval takes since the code book is divided into two or more steps and the size of the code book per step is reduced by for example, the B/L bit (B is the whole number of bits and L is a number of

stages here) by this approach is $L \times 2B/L$. It decreases. Moreover, memory space required for code book storing is reduced similarly. However, by this approach, since it is learning and looking for the code book of each stage independently, compared with one step of B bit, the engine performance has the trouble of falling greatly.

[0005] The purpose of this invention solves the trouble mentioned above, and is by comparatively small amount of operations and memory space to offer the good voice coding method of tone quality with a low bit rate, especially the bit rate below 4.8kbs/s.

[0006]

[Means for Solving the Problem] Divide the method of this invention into a frame for every timing which defined the sound signal to input beforehand, and it carries out a fragmentation rate to further two or more subframes. The spectrum parameter count section which computes the spectrum parameter with which the spectrum-description of said sound signal is expressed to said at least one subframe, The spectrum parameter quantization section which quantizes the spectrum parameter of said subframe of the location specified beforehand using a quantization code book, The mode classification section which calculates the characteristic quantity as which said sound signal was determined beforehand, and classifies the sound signal of said frame into one of two or more kinds of modes, The weighting section which searches for the weighting signal which gave audibility weighting to said sound signal according to said spectrum parameter obtained in said spectrum parameter count section, The mode classification result in said mode classification section, the spectrum parameter obtained in said spectrum parameter count section, the spectrum parameter quantized in said spectrum parameter quantization section, and said weighting signal are embraced. The adaptation code book section sent out in quest of the parameter showing the pitch of said sound signal corresponding to the mode, It responds to said weighting signal, the sending-out parameter of said adaptation code book section, said spectrum parameter, and said quantized spectrum parameter. It has the sound-source quantization section which emits the output signal which looked for the sound-source code book and gain code book which consist of two or more steps, and quantized the excitation signal of said sound signal.

[0007]

[Function] An operation of the voice coding method by this invention is shown.

[0008] A sound signal is divided into a frame (for example, 40ms), and it divides into a subframe (for example, 8ms) further. In the spectrum parameter count section, from at least one subframe (the 1st, 3rd, and 5th subframe for example, among five subframes), well-known LPC analysis is performed and it asks for a spectrum parameter (LPC parameter). In the spectrum parameter quantization section, the LPC parameter corresponding to the subframe (for example, the 5th subframe) defined beforehand is quantized using a quantization code book. Here, as a code book, both a vector quantization code book a scalar quantity child-ized code book and a vector-scalar quantity child-ized code book can be used.

[0009] Next, the characteristic quantity beforehand defined from the sound signal of a frame is calculated, the threshold beforehand determined as this value is compared, and it classifies into two or more kinds (for example, four kinds) of modes for every frame. Next, in the audibility weighting section, an audibility weighting signal is calculated for every subframe by the bottom type (1) using the spectrum parameter a_i ($i = 1$ thru/or P) of the 1st, 2nd, and 5th subframe. however -- the spectrum parameter of the 2nd and 4th subframes -- respectively -- the [for example, / the 1st and 3rd subframe and] -- it is obtained by carrying out linear interpolation of the spectrum parameter of the 3 5th subframe.

[0010]

$$X_w(z) = x(z) \cdot \left[1 - \frac{\sum_{i=1}^P a_i z^{-i}}{1 - \sum_{i=1}^P a_i \gamma^i z^{-i}} \right] \quad \dots (1)$$

[0011] Here, $x(z)$ and $X_w(z)$ are z-transformation of the sound signal of a frame, and an audibility weighting signal, respectively. P is the degree of a spectrum parameter. Moreover, gamma is a constant for controlling the amount of audibility weighting, and is usually chosen about as 0.8.

[0012] Next, in the adaptation code book section, Delay T and Gain β are calculated as a parameter about a pitch for every subframe to an audibility weighting signal. Delay corresponds to a pitch period here. Refer to said reference 2 for the calculus of the parameter of an adaptation code book. Moreover, in order to improve the engine performance of an adaptation code book especially to a female speaker, the delay for every subframe can also be expressed with the fractional value instead of an integral value for every sampling time. "pitch pre DIKUTAZU specifically according to KURUN (P. Kroon) and Atar (B. Atal), and Wiz - yes, 661 thru/or a 664-page paper (reference 4), etc. can be referred to in minutes (Proc.ICASSP) 1990 of the international congress about - temporal

resolution (Pitch predictors with high temporal resolution)", sound, voice, and signal processing. Although that was required 7 bits will increase to about 8 bits by making it a fractional value if the amount of delay for every subframe is expressed with an integral value by doing in this way, about a female sound, tone quality is improved remarkably.

[0013] Furthermore, in order to reduction-ize the amount of operations about count of the parameter of an adaptation code book, it asks for two or more kinds of candidates of delay by open-loop retrieval for every subframe from an audibility weighting signal sequentially from what maximizes a bottom type (2) first.

[0014]

$$D(T) = P2(T)/Q(T) \quad (2)$$

However, [0015]

$$P(T) = \sum_{n=0}^{N-1} x_w(n) x_w(n-T) \quad \dots(3)$$

$$Q(T) = \sum_{n=0}^{N-1} x_w(n-T)^2 \quad \dots(4)$$

[0016] It comes out. For every subframe, by open-loop retrieval, it asks for at least one kind of delay candidate, and after that, for every subframe, it searches near said candidate and a pitch period (delay) and gain are searched for by the closed loop retrieval using the drive excitation signal of the past frame by the above. see Japanese Patent Application No. No. (reference 5) 103262 [three to] etc. about a concrete approach, for example)

The amount of delay of an adaptation code book can reduction-ize sharply transmission amount of information required in order to transmit delay of an adaptation code book in the voiced section compared with the approach of transmitting the amount of delay independently for every subframe, by taking the difference of the amount of delay between subframes, and transmitting difference between subframes, since correlation is very high. For example, if the difference of the amount of delay with a just before subframe is transmitted by the triplet by the 2-5th subframes for every frame by expressing the amount of delay with a fractional value, and transmitting by 8 bits by the 1st subframe, compared with the case where 8 bits is transmitted, transmission amount of information can be reduction-ized from 40 bits per frame to 20 bits by all subframes.

[0017] Next, the sound-source quantization section is searched for the sound-source code book which consists of two or more steps of code books for vector quantization, and a code vector is chosen for every stage so that the error power of the above-mentioned weighting signal and the signal in which weighting playback was carried out by each code vector in a sound-source code book may be minimized. For example, if the sound-source code book consists of two steps of code books, retrieval of a code vector will be performed according to a bottom type (5).

[0018]

$$D = \sum_{n=0}^{N-1} [x_w(n) - \beta v(n-T) * h_w(n) - \gamma_1 c_{1j}(n) * h_w(n) - \gamma_2 c_{2i}(n) * h_w(n)]^2 \quad \dots(5)$$

[0019] In an upper type, $\beta v(n-T)$ is the adaptation code vector calculated in closed loop retrieval of the adaptation code book section, and β is the gain of an adaptation code vector. $c_{1j}(n)$ and $c_{2i}(n)$ express the j -th code vector of the 1st step and the 2nd step of code book, and the i -th code vector, respectively. Moreover, $h_w(n)$ is an impulse response showing the property of a bottom-type (6) weighting filter. Moreover, γ_1 and γ_2 are the optimal gain about the 1st step and the 2nd step of code book, respectively.

[0020]

$$H_w(z) = [1 - \sum_{i=1}^p a_i z^{-i} / 1 - \sum_{i=1}^p a_i \gamma_1' z^{-i}] [1 / 1 - \sum_{i=1}^p a_i \gamma_2' z^{-i}] \quad \dots(6)$$

[0021] Here, γ is a constant which controls audibility weighting of a formula (1).

[0022] Next, after searching for the code vector which minimizes the formula (5) of a sound-source code book, it looks for a gain code book so that a bottom type (7) may be minimized.

[0023]

$$D = \sum_{n=0}^{N-1} [x_w(n) - \beta v(n-T) * h_w(n) - \gamma_{1k} c_{1k}(n) * h_w(n) - \gamma_{2k} c_{2k}(n) * h_w(n)]^2 \quad \dots(7)$$

[0024] gamma1 k and gamma2k show the k-th gain code vector of a two-dimensional gain code book here.

[0025] In order to reduction-ize the amount of operations when searching for the optimal code vector of a sound-source code book The candidate (the 1st step is m1 class and the 2nd step is m2 class) of two or more kinds of sound-source code vectors may be chosen for every stage, it may search for the 2nd step of candidate's total combination (m1 xm2) with the 1st step after selection, and the combination of the candidate who makes a formula (5) min may be chosen.

[0026] Moreover, when looking for a gain code book, it looks for a gain code book according to a formula (7), and you may make it search for the combination of a gain code vector which minimizes error power, and a sound-source code vector from the combination of the candidate who chose only the number beforehand set to the small order of error power among all the combination of the candidate of the above-mentioned sound-source code vector, or combination. Although the amount of operations will increase if it does in this way, the engine performance improves.

[0027] Next, in the mode classification section of the embodiment indicated to claim 2 of the claim of this invention, accumulation pitch prediction distortion is used as characteristic quantity. First, the pitch period candidate T chosen by open-loop retrieval of the adaptation code book section for every subframe is asked for pitch prediction error distortion for every subframe as pitch prediction distortion according to a bottom type (8).

[0028]

$$D_1 = \sum_{n=0}^{N-1} x_{w1}^2(n) - P_1^2(T) / Q_1(T) \quad \dots(8)$$

[0029] 1 is a subframe number here. And the accumulation prediction error power of the whole frame is found by the bottom type (9), the threshold beforehand determined as this value is compared, and it classifies into two or more kinds of modes.

[0030]

$$D = 1/M \sum_{i=1}^M D_i \quad \dots(9)$$

[0031] For example, supposing it forms four kinds of modes, three kinds of thresholds will be prepared and a mode classification will be performed for the value of a formula (9) as compared with three kinds of thresholds. In addition, as pitch prediction distortion, pitch prediction gain etc. can also be used in addition to the above.

[0032] To some of modes classified according to the mode classification section, to a training signal, the spectrum quantization code book is created beforehand, and when encoding, in the spectrum parameter quantization section of the embodiment of this invention indicated to claim 3, it is used using mode information, changing a spectrum quantization code book. Although only the class to change increases, in the whole sum total, the memory space which stores a code book if it does in this way can become equivalent to having the code book of bigger size, and can raise the engine performance, without increasing transmission amount of information.

[0033] In the sound-source quantization section of the embodiment of this invention indicated to claim 4, a training signal is beforehand classified for every mode, and a sound-source code book different the whole mode defined beforehand and the gain code book are created, and when encoding, it is used using mode information, changing a sound-source code book and a gain code book. Although only the class to change increases, in the whole sum total, the memory space which stores a code book if it does in this way can become equivalent to having the code book of bigger size, and can raise the engine performance, without increasing transmission amount of information.

[0034] In the sound-source quantization section of the embodiment of this invention furthermore indicated to claim 5, at least one of two or more steps of code books has regular pulse composition of the rate of infanticide (for example, the rate of infanticide = 2) as which the element of a code vector was determined beforehand. Here, it becomes rate = of infanticide 1, then the usual configuration. By making it such a configuration, the amount of memory required for storing of a sound-source code book can be reduced to the rate of 1-/infanticide (for example, the rate of infanticide = if 2 1/2). Moreover, -izing also of the amount of operations required for sound-source code

book retrieval can be mostly carried out [****] to below the rate of 1-/infanticide. Furthermore, by thinning out and pulse-izing the element of a sound-source code vector, especially, in the audio vowel section, since an important pitch pulse can be expressed more to fitness on audibility, tone quality improves.

[0035]

[Example] Next, this invention is explained with reference to a drawing.

[0036] Drawing 1 is the block diagram showing the 1st example of this invention. In this drawing, the sound signal inputted from an input terminal 100 is divided into every frame (for example, 40ms) in the frame dividing network 110, and is divided into a subframe (for example, 8ms) still shorter than a frame in the subframe dividing network 120.

[0037] In the spectrum parameter count circuit 200, it calculates in the degree (for example, degree $P=10$) which the aperture (for example, 24ms) longer than subframe length was applied, started voice to the sound signal of at least one subframe, and was able to define the spectra parameter beforehand. Since it changes a lot in time in the transition segment between a consonant and a vowel and the amount of operations required for analysis will increase if it is made such although to analyze for every short time amount is more desirable, especially a spectrum parameter is made to calculate a spectrum parameter to any L subframes ($L>1$) in a frame (for example, referred to as $L=3$ the 1st, 3, and 5 subframe). And in the subframe (here the 2nd and 4 subframe) which was not analyzed, what carried out linear interpolation of the spectrum parameter of the 1st, the 3rd subframe, and the 3rd and the 5th subframe on below-mentioned LSP is used as a spectrum parameter, respectively. Well-known LPC analysis, the Berg (Burg) analysis, etc. can be used for count of a spectrum parameter here. Burg analysis is used in this example. The detail of Burg analysis is indicated by 82 of the book (Corona Publishing, 1988 annual publications) entitled "the signal analysis and system identification" by the inside slot, for example thru/or 87 pages (reference 6).

[0038] further -- the spectrum parameter count circuit 200 -- Burg -- linear-predictor-coefficients α_{pi} ($i=1$ thru/or 10) calculated by law is changed into the line spectrum pair (LSP) parameter suitable for quantization or interpolation. Here, the transformation method from linear predictor coefficients to LSP uses the paper (the Institute of Electronics and Communication Engineers paper magazine, J64-A, 599 or 606 pages, 1981) (reference 7) it was ["speech information compression by the line spectrum pair (LSP) voice-analysis composite system" by Sugamura and others] entitled. that is, the 1st, 3, and 5 subframe -- Burg -- the linear predictor coefficients for which it asked by law are changed into an LSP parameter, and LSP of the 2nd and 4 subframe is calculated by linear interpolation, and inverse transformation of LSP of the 2nd and 4 subframe is carried out, it returns to linear predictor coefficients, and linear-predictor-coefficients α_{pi} ($i=1$ thru/or 10, $i=1$, or 5) of the 1st thru/or 5 subframes is outputted to the audibility weighting circuit 230. Moreover, LSP of the 1st thru/or the 5th subframe is outputted to the spectrum parameter quantization circuit 210.

[0039] In the spectrum parameter quantization circuit 210, the LSP parameter of the subframe defined beforehand is quantized efficiently. The LSP parameter of the 5th subframe is quantized in this example, using vector quantization as a quantizing method. The technique of vector quantization of an LSP parameter can use the well-known technique. In the spectrum parameter quantization section 210, the 1st thru/or the LSP parameter of the 4th subframe are further restored based on the LSP parameter quantized by the 5th subframe (see Japanese Patent Application No. No. (reference 8) 297600 [two to], Japanese Patent Application No. No. (reference 9) 261925 [three to], Japanese Patent Application No. No. (reference 10) 155049 [three to], etc.). In this example, linear interpolation of the quantization LSP of the 5th subframe of the frame of one past is carried out to the quantization LSP parameter of the 5th subframe of the present frame, and LSP of the 1st thru/or the 4th subframe is restored. That is, after choosing one kind of code vector which minimizes the error power of LSP before quantization, and LSP after quantization, linear interpolation can restore LSP of the 1st thru/or the 4th subframe. What is necessary is to evaluate the accumulation distortion by the bottom type (10) about each candidate, and just to make it choose the group of Interpolation LSP with the candidate who minimizes accumulation distortion, after making two or more candidate selection of the code vector which minimizes said error power, in order to raise the engine performance furthermore.

[0040]

$$D = \sum_{i=1}^5 \sum_{j=1}^{10} c_{ij} b_{ij} [1sp_{ij} - 1sp'_{ij}]^2 \quad \dots (10)$$

[0041] Here, they are $1sp_{ij}$ and $1sp'_{ij}$. They are LSP before quantization of the 1st subframe, and LSP of the 1st

subframe restored after quantization, respectively. Moreover, bil is the weighting factor for which it asked with the application of the bottom type (11) from LSP before quantization of the 1st subframe.

[0042]

$$\text{bil} = (1/[1\text{spi}, l - 1\text{spi} - 1, l]) + (1/[1\text{spi} + 1, l - 1\text{spi}, l]) \quad (11)$$

Moreover, ci It is a weighting factor to the direction of a degree of LSP, for example, can ask using a bottom type (12).

[0043]

$$c_i = -1.0 \text{ (} i = 1 \text{ thru/or } 8 \text{) and } 0.8 \text{ (} i = 9 \text{ thru/or } 10 \text{)} \quad (12)$$

LSP of the 1st restored by the above thru/or the 4th subframe and the quantization LSP of the 5th subframe are changed into linear-predictor-coefficients alpha'il (i= 1 thru/or 10, l= 1, or 5) for every subframe, and it outputs to the impulse response count circuit 310. Moreover, the index showing the code vector of the quantization LSP of the 5th subframe is outputted to a multiplexer 400.

[0044] You may make it choose the group of the code vector which prepares by the number of bits (for example, 2 bits), restores LSP of the 1st thru/or the 4th subframe to each of these patterns, evaluates a formula (10), and minimizes a formula (10), and a interpolation pattern which was able to define the storage pattern of LSP beforehand instead of linear interpolation in the above-mentioned actuation. Although transmission information will increase only the number of bits of a storage pattern if it does in this way, a time change within the frame of LSP can be expressed more to a precision. Here, using the LSP data for training, a storage pattern may be learned beforehand, and may be created, and the pattern defined beforehand may be stored.

[0045] In the mode classification circuit 245, the prediction error power of a spectrum parameter is used as characteristic quantity for performing a mode classification. The linear predictor coefficients calculated by the spectrum parameter count circuit 200 are inputted by five subframes, it changes into K parameter, and the accumulation prediction error power E for five subframes is calculated by the bottom type (13).

[0046]

$$E = 1/5 \sum_{l=1}^5 G_l \quad \dots (13)$$

[0047] However, [0048]

$$G_l = P_l \cdot \left(\prod_{i=1}^{10} [1 - k_{il}^2] \right) \quad \dots (14)$$

[0049] Come out, and it is and is P1. It is PAWA of the input signal of the 1st subframe. Next, the value of E is classified into two or more kinds of modes as compared with the threshold which was able to be defined beforehand. For example, when classifying into four kinds of modes, it carries out by comparing with three kinds of thresholds. While the mode information classified and acquired is outputted to the adaptation code book circuit 300, the index (it is 2 bits at the time of four kinds of mode information) showing mode information is outputted to a multiplexer 400.

[0050] From the spectrum parameter count circuit 200, the weighting circuit 230 inputs linear-predictor-coefficients alpha'il (i= 1 thru/or 10, l= 1, or 5) for every subframe, performs audibility weighting to the sound signal of a subframe based on a formula (1), and outputs an audibility weighting signal.

[0051] The reply signal count circuit 240 inputs linear-predictor-coefficients alpha'il for every subframe from the spectrum parameter count circuit 200, using the value of the filter memory which inputs linear-predictor-coefficients alpha'il interpolated [was quantized and] and restored for every subframe, and is saved from the spectrum parameter quantization circuit 210, calculates the reply signal set to input signal d(n)=0 by one subframe, and outputs it to a subtractor 250. Here, it is a reply signal xz. (n) is expressed with a bottom type (15).

[0052]

$$x_2(n) = d(n) - \sum_{i=1}^{10} \alpha_i \cdot d(n-i) + \sum_{i=1}^{10} \alpha_i \gamma^i \cdot y(n-i) + \sum_{i=1}^{10} \alpha_i' \gamma^i \cdot x_z(n-i) \quad \dots (15)$$

[0053] Here, gamma is the same value as the case of a formula (1).

[0054] By the bottom type, a subtractor 250 subtracts a reply signal from an audibility weighting signal by one subframe, and outputs $xw'(n)$ to the adaptation code book circuit 300.

[0055]

$$xw'(n) = xw(n) - xz(n) \quad (16)$$

The impulse response count circuit 310 is the impulse response hw of the weighting filter by which z -transformation is expressed with a bottom type. Only the mark L which were able to define (n) beforehand are calculated and it outputs to the adaptation code book circuit 300 and the sound-source quantization circuit 350.

[0056]

$$H_w(z) = [1 - \sum_{i=1}^{10} \alpha_i z^{-i} / 1 - \sum_{i=1}^{10} \alpha_i \gamma^i z^{-i}] \cdot [1 / 1 - \sum_{i=1}^{10} \alpha_i \gamma^i z^{-i}] \quad \cdots (17)$$

[0057] The adaptation code book circuit 300 inputs the mode information from a mode classification circuit, and asks for a pitch parameter only at the time of the mode defined beforehand. Here, supposing the mode is large as the threshold at the time of those with 4 modes and a mode classification becomes the mode 3 from the mode 0, since the mode 0 is considered to correspond to the vowel section for child Otobe, the mode 1, or 3, a pitch parameter will be made to ask for the adaptation code book circuit 300 only to the mode 1 thru/or 3. First, in open-loop retrieval, two or more kind (for example, M kinds) choice of the integer delay candidate who maximizes a formula (2) for every subframe is made to the output signal of the audibility weighting circuit 230. Furthermore, in the short field of delay, it asks for two or more kinds of decimal delay candidates [near the integer delay] using the technique of said reference 4 grade from a candidate's each (delay is 20 thru/or 80), and at least one kind of decimal delay candidate who finally maximizes said formula (2) is chosen for every subframe. Below, since it is easy, the number of candidates is made into one kind, and sets to dl ($l = 1$ thru/or 5) delay chosen one kind for every subframe. Next, it sets to closed loop retrieval and is dl for every subframe based on drive excitation signal [of the past frame] $v(n)$. Index ld which evaluates a bottom type (18) to several point epsilon of near appointed beforehand, asks for the delay which maximizes the value for every subframe, and expresses delay ld outputs to a multiplexer. Reference 5 grade can be referred to about the detail of heuristics. Moreover, an adaptation code vector is calculated by the bottom type (21), and it outputs to the sound-source quantization circuit 350.

[0058]

$$D'(dl + \epsilon) = P'^2 / (dl + \epsilon) \quad Q(dl + \epsilon) \quad (18)$$

However, [0059]

$$P'(dl + \epsilon) = \sum_{n=0}^{N-1} xw'(n) [v(n - (dl + \epsilon)) * h_w(n)] \quad \cdots (19)$$

$$Q(dl + \epsilon) = \sum_{n=0}^{N-1} [v(n - (dl + \epsilon)) * h_w(n)]^2 \quad \cdots (20)$$

[0060] Here, it is $h_w(n)$ is the output of an impulse response count circuit.

[0061]

$$q(n) = \beta - v(n - (dl + \epsilon)) * h_w(n) \quad (21)$$

$$\text{However, } \beta = P'(dl + \epsilon) / Q(dl + \epsilon) \quad (22)$$

Moreover, as the term of an operation explained, in the voiced section (for example, the mode 1 thru/or 3), the difference of delay can be taken between subframes and difference can also be transmitted. In such a configuration (for example, the 1st subframe of a frame), 8 bit transmissions can be carried out by decimal delay, and the difference [subframe / front] of delay can be transmitted per subframe by the 2-5th subframes at a triplet. Moreover, at the time of open-loop delay retrieval, by the 2-5th subframes, it supposes that it searches for the near value of delay of a front frame by the triplet, and the candidate of delay is not further chosen for every subframe, but the error power of accumulation is found by five subframes from the pass for five subframes of a delay candidate, and it asks for the pass of the delay candidate who minimizes this, and outputs to closed loop retrieval. Closed loop retrieval is searched near the delay value acquired by closed loop retrieval by the front subframe by the triplet, and the index corresponding to the delay value for every subframe which calculated and calculated the final

delay value is outputted to a multiplexer 400 by it.

[0062] The sound-source quantization circuit 350 inputs the output signal of a subtractor 250, the output signal of the adaptation code book circuit 300, and the output signal of the impulse response count circuit 310, and looks for the vector quantization code book which consists of two or more steps first. At drawing 1, it is 3511-351 Ns of sound-source code books about two or more kinds of vector quantization code books. It expresses by carrying out. Below, since it is easy, a number of stages is set to 2. Retrieval of the code vector of each stage follows the bottom type (23) which corrected the formula (5).

[0063]

$$D = \sum_{n=0}^{N-1} [x'w(n) - \beta \cdot q(n) - \gamma_{1c1}(n) * h_w(n) - \gamma_{2c2}(n) * h_w(n)]^2 \quad \dots (23)$$

[0064] However, $x'w(n)$ is the output signal of a subtractor 250. In addition, in order not to use an adaptation code book, it searches for the code vector which minimizes a bottom type (24) in the mode 0 instead of a formula (23).

[0065]

$$D = \sum_{n=0}^{N-1} [x'w(n) - \gamma_{1c1}(n) * h_w(n) - \gamma_{2c2}(n) * h_w(n)]^2 \quad \dots (24)$$

[0066] Although there is various heuristics of the 1st step for minimizing a formula (23) and the 2nd step of code vector, the 1st step and the 2nd step to two or more kinds of candidates are chosen here, after that, combination retrieval of both candidates is performed and the combination of the candidate who minimizes distortion of a formula (23) is determined. Moreover, vector quantization [the 1st step and the 2nd step of] code book is beforehand designed in consideration of the above-mentioned heuristics using a lot of voice databases. The indexes $Ic1$ and $Ic2$ of the 1st step and the 2nd step of code vector determined by the above are outputted.

[0067] Moreover, the sound-source quantization circuit 350 also performs retrieval of a gain code book. The gain code book is expressed with drawing 1 as 355. The mode 1 which uses an adaptation code book thru/or 3 are searched for a gain code book so that a bottom type (25) may be minimized using the index as which the sound-source code book was determined.

[0068]

$$D_k = \sum_{n=0}^{N-1} [x'w(n) - \beta' \cdot q(n) - \gamma'_{1k} c_{1k}(n) * h_w(n) - \gamma'_{2k} c_{2k}(n) * h_w(n)]^2 \quad \dots (25)$$

[0069] Here, suppose that the gain of the gain of an adaptation code vector, the 1st step, and the 2nd step of sound-source code vector is quantized using the gain code book of a three dimension. Here, (β_k , γ_{1k} , γ_{2k}) are the k -th code vector. In order to minimize a formula (25), you may ask for the gain code vector which minimizes a formula (25) to all gain code vectors ($k=0$ thru/or $2B-1$), two or more kind preliminary selection of the candidate of a gain code vector is made, and what minimizes a formula (25) may be chosen from two or more of the kinds. Index Ig which shows the selected gain code vector after gain code vector decision It outputs. On the other hand, it looks for a gain code book in the mode in which an adaptation code book is not used so that a bottom type (26) may be minimized. Here, a two-dimensional gain code book is used.

[0070]

$$D_k = \sum_{n=0}^{N-1} [x'w(n) - \gamma'_{1k} c_{1k}(n) * h_w(n) - \gamma'_{2k} c_{2k}(n) * h_w(n)]^2 \quad \dots (26)$$

[0071] The weighting signal count circuit 360 inputs the output parameter and each index of a spectrum parameter count circuit, reads the code vector corresponding to it from an index, and asks for drive excitation signal $v(n)$ based on a bottom type first.

[0072]

$$v(n) = \beta' v(n-d) + \gamma'_{1k} c_{1k}(n) + \gamma'_{2k} c_{2k}(n) \quad (27)$$

However, it is referred to as $\beta'=0$ in the mode which does not use an adaptation code book. Next, the output

parameter of the spectrum parameter count circuit 200 and the output parameter of the spectrum parameter quantization circuit 210 are used, and it is the weighting signal sw by the bottom type (28). (n) is calculated for every subframe and it outputs to the reply signal count circuit 240.

[0073]

$$S_w(n) = v(n) - \sum_{i=0}^{10} a_i v(n-i) + \sum_{i=0}^{10} a_i \gamma^i p(n-i) + \sum_{i=0}^{10} a_i' \gamma^i sw(n-i) \quad \dots(28)$$

[0074] Explanation of the 1st example of this invention is finished by the above.

[0075] Drawing 2 is the block diagram showing the 2nd example of this invention. Since the component to which this example gave the same reference mark as the example of drawing 1 about the mode classification circuit 410 performs the same actuation as the case of drawing 1, it omits explanation.

[0076] Based on the above-mentioned formula (2) and (3), for every subframe, the open-loop count circuit 421 in the adaptation code book circuit 420 asks for at least one kind of candidate of delay, and outputs him to the closed loop count circuit 422. Furthermore, the pitch prediction error power of a formula (29) is calculated for every subframe.

[0077]

$$P_{G1} = \sum_{n=0}^{N-1} x_{w1}^2(n) - P_1^2(T) / Q_1(T) \quad \dots(29)$$

[0078] And P_{G1} is outputted to the mode classification circuit 410.

[0079] From mode information and the open-loop count circuit 421, the closed loop count circuit 422 inputs at least one kind of delay candidate, and an audibility weighting signal for every subframe, and performs the same actuation as the closed loop retrieval section in the adaptation code book circuit 300 of the 1st example.

[0080] The mode classification circuit 410 follows a bottom type (30), and is the accumulation pitch prediction error power E_G as characteristic quantity. It asks, the mode is classified for this as compared with two or more kinds of thresholds, and mode information is outputted.

[0081]

$$E_G = 1 / 5 \sum_{i=1}^5 P_{G1} \quad \dots(30)$$

[0082] Above, explanation of the 2nd example is ended.

[0083] Drawing 3 is the block diagram showing the 3rd example of this invention. In drawing 3, since the component which attached the same reference mark as drawing 1 performs the same actuation as drawing 1, explanation is omitted. Setting to drawing 3, the spectrum parameter quantization circuit 450 is two or more kinds of quantization code book 4510 -451M-1 for spectrum parameter quantization. It is 4510 -451M-1 for every mode which has, inputted mode information from the mode information classification circuit 250, and was defined beforehand. It is used changing.

[0084] 4511-451Ns of quantization code books What is necessary is to classify a lot of spectrum parameters for training into the mode beforehand, and just to design the quantization code book for every mode defined beforehand. Since it becomes almost equivalent to code book size having increased several times, keeping the transmission amount of information of the index of a quantization spectrum parameter, and the amount of operations of code book retrieval the same as that of drawing 1 by taking such a configuration, the engine performance of spectrum parameter quantization is sharply improvable.

[0085] Explanation of the 3rd example is ended by the above.

[0086] Drawing 4 is the block diagram showing the 4th example of this invention. In drawing 4, since the component which attached the same reference mark as drawing 1 performs the same actuation as drawing 1, explanation is omitted. drawing 4 -- setting -- the sound-source quantization circuit 470 -- the vector quantization code books 47110-4711N of N stage ($N>1$) -- M set ($M>1$) and the gain code book 481 -- 4810 -481M-1 up to -- it

has M sets and has the vector quantization code book of 47110-471NM-1 (a total of NxM kinds). Using the mode information from the mode classification circuit 250, the vector quantization code book of N stage in the j-th set with which it was beforehand set of the M sets at the time of the mode defined beforehand is chosen, the gain code book in the j-th set defined beforehand is chosen, and an excitation signal is quantized.

[0087] What is necessary is to classify a lot of voice databases for every mode beforehand, and just to design a code book for every mode defined beforehand using the above-mentioned approach, when designing a sound-source code book and a gain code book. Since it becomes almost equivalent to code book size having increased M times by these, keeping the transmission amount of information of the index of a sound-source code book and a gain code book, and the amount of operations of sound present code book retrieval the same as that of drawing 1, the engine performance of sound-source quantization is sharply improvable.

[0088] It sets in the sound-source quantization circuit 350 of drawing 4, and is 3511-351 Ns of code books of N stage. It has and at least one of steps [them] has regular pulse composition of the rate of infanticide defined beforehand as shown in drawing 5. Drawing 5 shows the example of the rate $m=2$ of infanticide. If it is made a regular pulse configuration, since data processing is unnecessary in the location of zero, amplitude can reduction-ize the amount of operations required for code book retrieval to about $1/m$. Moreover, since amplitude does not need to store the amount of memory required to store a code book in the location of zero, either, -izing can be carried out [****] to about $1/m$. About the detail of a regular pulse configuration "A and the 6kbps. regular pulse, carcinoembryonic antigen rupee coder forehand Mobile Radio Communications () of Dell Prato and others (M. Delprat) [A 6kbps regular pulse CELP] coder communications" for (edited by Atar (Atal) --) mobile radio Kluwer Academic Since the paper (reference 11) it was [179 thru/or 188 pages] entitled can be referred to in Publishers and 1990, explanation is omitted here. By the above-mentioned approach, the code book of a regular pulse configuration is also learned beforehand.

[0089] Furthermore, when expressing the amplitude pattern of a different phase as a common pattern, designing a code book, and using only a phase at the time of coding, shifting in time, -izing of the amount of memory and the amount of operations can be carried out [****] to a pan one half at the time of $m=2$.

[0090] Moreover, in order to reduce the amount of memory, the configuration of a multi-pulse can also be taken besides a regular pulse configuration.

[0091] Explanation of the 4th example of this invention is ended by the above.

[0092] Various deformation is possible besides the example mentioned above, without spoiling the intention of this invention.

[0093] First, a spectrum parameter can use a parameter [**** / others] besides LSP.

[0094] When calculating a spectrum parameter by at least one subframe in a frame, change of RMS of a front subframe and a current subframe or change of PAWA is measured, and you may make it these change calculate a spectrum parameter to two or more big subframes in the spectrum parameter count circuit 200. If it does in this way, even if it reduces the number of subframes which will surely analyze a spectrum parameter and is analyzed, degradation of the engine performance can be prevented at an audio changing point.

[0095] Approaches [****], such as vector quantization, formation of a scalar quantity child, and formation of a vector-scalar quantity child, can be used for quantization of a spectrum parameter.

[0096] An interval scale [**** / others] can be used for selection of the interpolation pattern in a spectrum parameter quantization circuit besides a formula (10). For example, a bottom type (31) can also be used.

[0097]

$$D = \sum_{l=1}^5 R_l \sum_{i=0}^{10} c_{li} [1sp_l - 1sp'_i]^2 \quad \dots (31)$$

$$R_l = RMS_l / \left[\sum_{i=1}^5 RMS_i \right] \quad \dots (32)$$

[0098] Here, it is RMS1. It is RMS or PAWA of one subframe.

[0099] Moreover, it sets in a sound-source quantization circuit, and is gain gamma 1 at formula (23) - (26). gamma

2 can also be made the same. At this time, a gain code book serves as two-dimensional gain in the mode in which an adaptation code book is used, and serves as 1-dimensional gain in the mode in which an adaptation code book is not used. Moreover, the number of stages of a sound-source code book or the number of bits of the sound-source code book of each stage, and the number of bits of a gain code book are also changeable for every mode. For example, the number of the modes 0 is three and the mode 1 thru/or 3 can also make them two steps.

[0100] Moreover, although the amount of memory will increase if the configuration of a sound-source code book changes the code book for which the 2nd step is searched according to the code vector which is made to correspond to the 1st step code vector, designs the 2nd step of code book at the time for example, of a two-step configuration, and was chosen in the 1st step, the engine performance improves further.

[0101] Moreover, a scale [**** / others] can also be used for the interval scale at the time of retrieval of a sound-source code book and study.

[0102] moreover, a gain code book -- the number of transmitted bits -- the whole -- several time size -- the code book of **** thing size is learned beforehand, and some fields of said code book are assigned as a use field for every mode defined beforehand, and when encoding, it can also be used according to the mode, being able to change a use field.

[0103] Moreover, in retrieval in a sound-source quantization circuit, it is an impulse response hw like formula (19) - (21) and type (23) - (26) to retrieval in an adaptation code book circuit, and a list, respectively. Although calculated by collapsing using (n), a filtering operation can also perform this using a weighting filter by which transfer characteristics are expressed with a formula (6). Although the amount of operations will increase if it does in this way, the engine performance improves further.

[0104]

[Effect of the Invention] As stated above, since voice was classified into the mode using audio characteristic quantity and the approach of the approach of quantization of a spectrum parameter, actuation of an adaptation code book, and sound-source quantization is changed with the mode, according to this invention, as compared with the conventional method, good tone quality is acquired also with a lower bit rate.

[Translation done.]

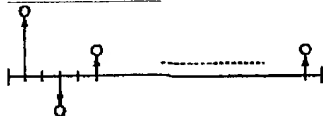
* NOTICES *

JPO and NCIPi are not responsible for any damages caused by the use of this translation.

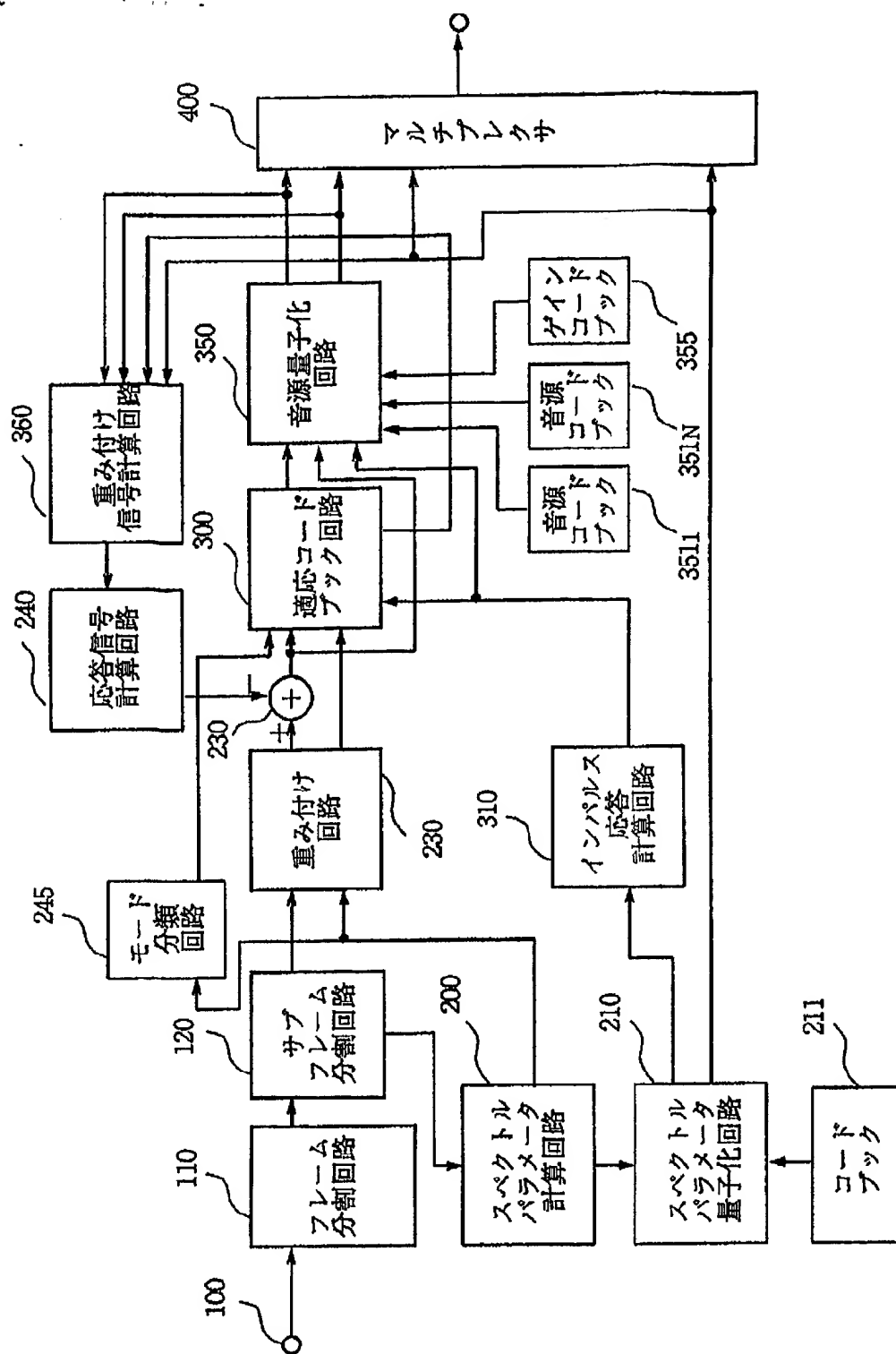
1. This document has been translated by computer. So the translation may not reflect the original precisely.
2. **** shows the word which can not be translated.
3. In the drawings, any words are not translated.

DRAWINGS

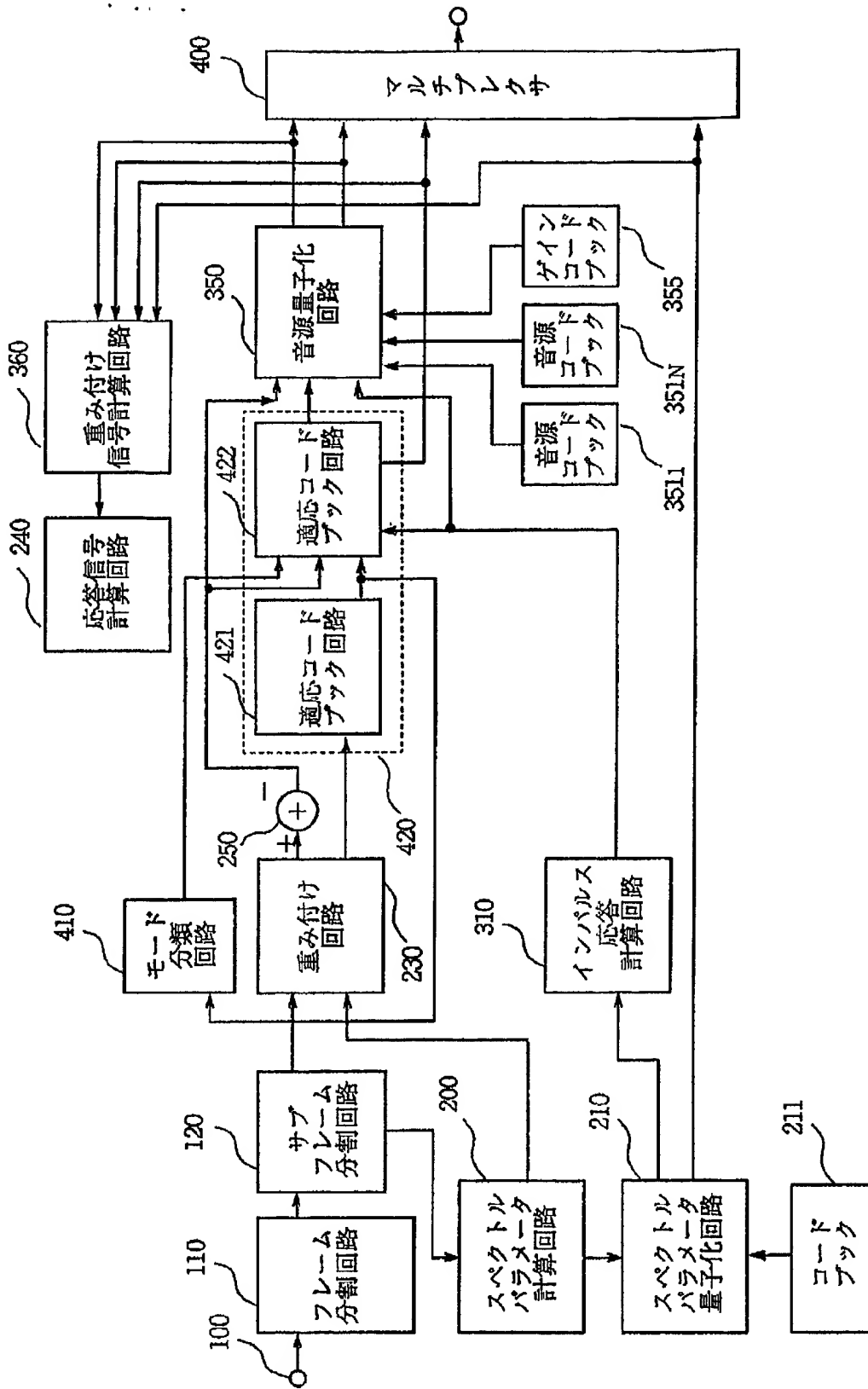
[Drawing 5]



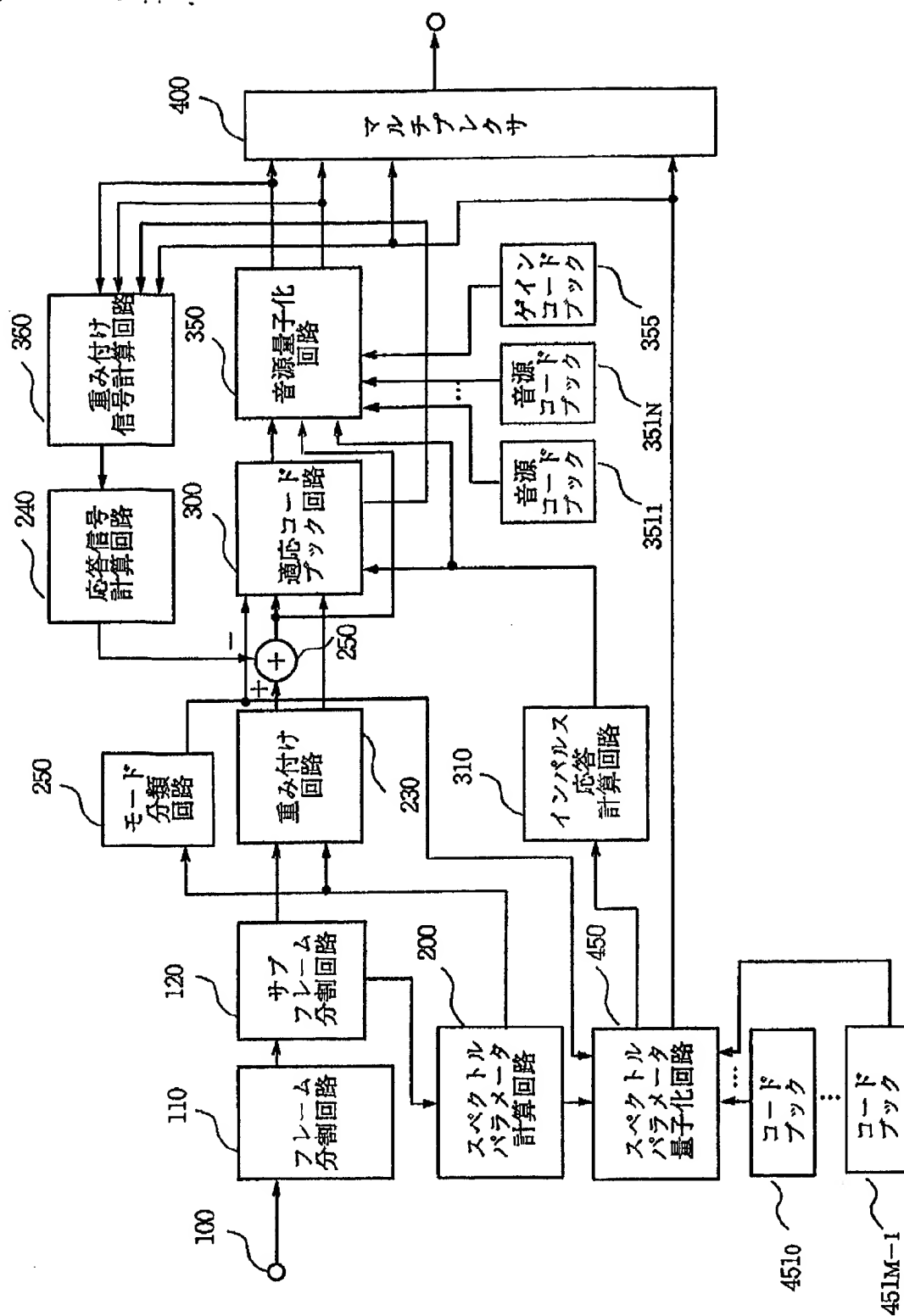
[Drawing 1]



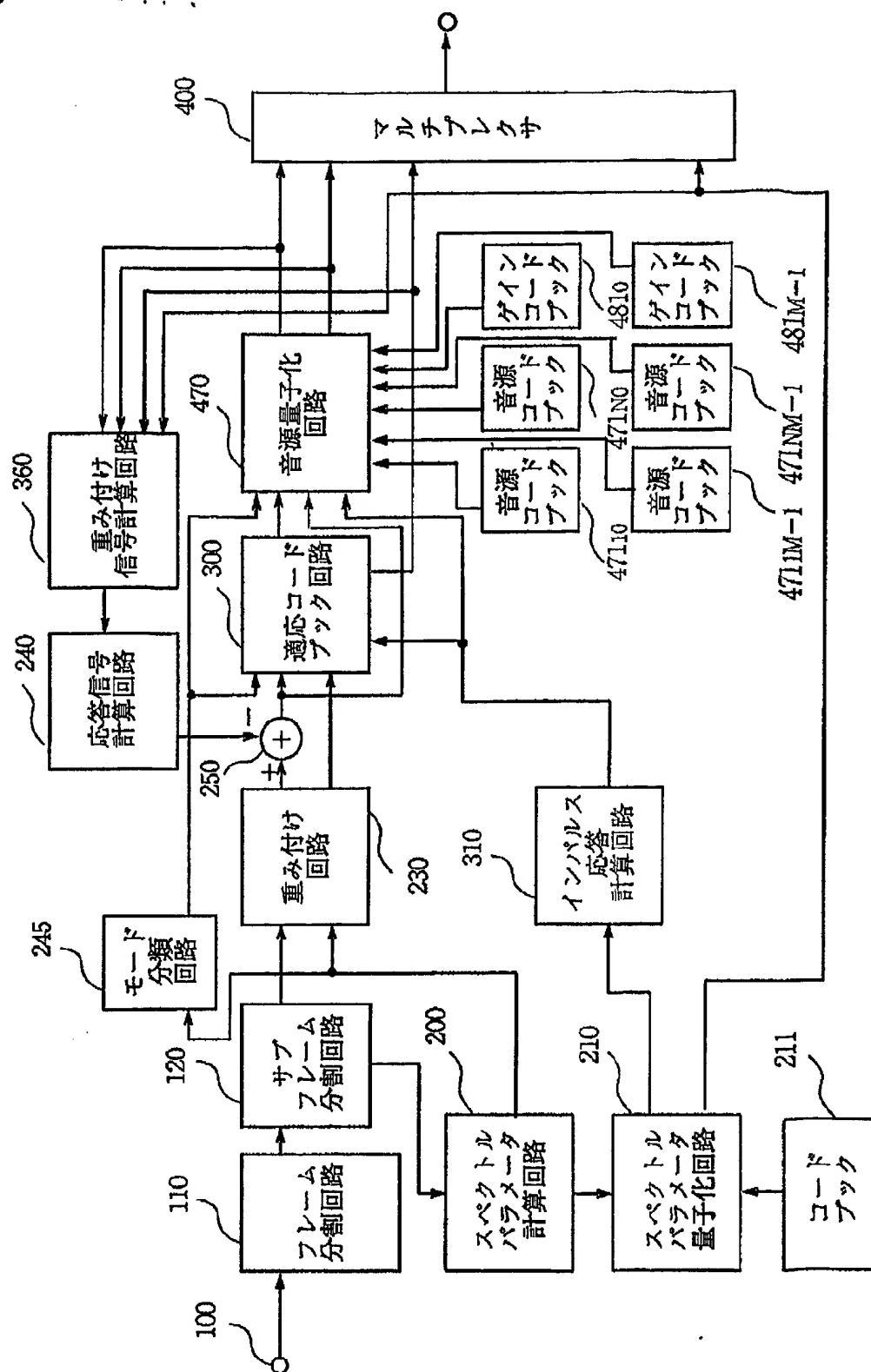
[Drawing 2]



[Drawing 3]



[Drawing 4]



[Translation done.]